Προτεινόμενα Θέματα Διπλωματικής Εργασίας

από Δρ. Στασινό Κωνσταντόπουλο, ΕΚΕΦΕ “Δημόκριτος”

A. Establishing a shared representation of world objects

Each AI system establishes its internal representation during training without making any reference to a prior representation or to the representation learned by other AI systems. This makes interfacing between AI systems or between an AI system and its user challenging, as they lack a shared conceptualization of the objects that make up their perception of the world. Although fine-tuning foundational models partially alleviates the AI-to-AI interaction challenge, the ability to align their internal representation with prior conceptualizations would give AI systems greater inspectactability and usability from the perspective of human users, but also greater flexibility to interoperate among themselves.

In this context, this thesis will explore how the internal, learned representation of an artificial visual perception system can be grounded in concepts of the physical world. where two AI agents try to establish a vocabulary, through repreated trial-and-feedback interactions, so that they can coordinate on a shared task.

An example of such an interaction can be based on how images are semantically segmented. The two agents must have been trained independently (without supervision or with weak supervision) to learn different models for semantic segmentation. As a consequence, their semantic categories are not perfectly aligned. The goal of the “describing agent” D is to convey to the “guessing agent” G the meaning of a specific object in a specific image, although D does not know the symbol (if any) that exactly matches this object. The agents receive feedback about how closely the segment guessed by G matches the segment that D wanted to convey, and iterate.

Within this broad setting, the student will select data and background methods (for semantic segmentation among others), design a specific communication mechanism, implement and execute the experiment, and analyse the results.

B. Situational awareness in the physical world

Semantic segmentation allows an AI system to abstract sensor data into the objects that are perceived and the class they belong to. *Data association* allows an AI system to track the senor-data segments of each object across frames so that objects maintain their identity as the scene evolves. *Spatial inference* allows an AI system to reason about the relative positions of objects and to reasonably complete its internal representation of the scene in the face of occlusions and uncertainties. Put together, these capabilities allow an AI system to be aware of its environment so that it can reliably plan how to proceed to accomplish its mission.

However, the integration of these capabilities into a unified situational awareness framework faces two challenges that reinforce each other: If seen as steps in a sensor data processing pipeline, they form a circular dependency because they all rely on the other two to manage uncertainty. But to tackle them as a single capability would mean to integrate machine vision methods for segmentation (where the best methods are machine learned models) with mathematical modelling for object tracking and with symbolic inference for spatial inference.

As an example, consider a robot learning about the world by observing a scene where a spoon is inserted into a (visually solid-looking) jar, scoops some oil (which would gain a distinct identity from the jar it came out off), and transfers it into a dish, at which point the oil looses its identity. But if spilled on the floor, the oil not only maintains its identity but our robot is well advised to remain aware of its presence and position and avoid it.

Within this broad setting, the student will explore integrating at least two of semantic segmentation, data association, and spatial inference. There are multiple integration approaches to experiment with. Indicatively, consider the joint optimization of two distinct components (one data-driven model and one knowledge-based inference engine) or automatic differentiation for training a network that combines data-driven and knowledge-based functions. The student is welcome to formulate alternative approaches, as long as they are neither purely data-driven nor purely knowledge-based. The student is expected to base their work on existing data and methods, and focus not on the methods themselves, but on how they can interoperate to mutually reduce each other's uncertainty.

C. Task planning in the digital world

Linked data is the Web-native way to publish structured data, emphasising the value that is added not only by semantically linking internally to the dataset, but also across datasets. The vision of linked data is for the Web to become the global database for data processing systems, besides its human users. But Linked Data is not as linked as it was envisaged. Syntactic heterogeneity is manageable, as connectors exist that can abstract over all the different network protocols and query languages. Semantic heterogeneity, on the other hand, is practically unmanageable as the rate at which data is published thwarts any manual curation effort.

Let us imagine an AI system that is aware of the major data portals and catalogues, foundational ontologies and linksets, dataset search engines, and, in general, of the resources that a human data analyst would be expected to be aware of. When tasked with an (at first glance) impossible data request, the AI system should autonomously search in data portals and search engines and reason about possible paths through linksets and (possibly) uncertain equivalences in order to fulfil the task. The systen would then present the user with a (possibly incomplete or uncertain) plan, highlighting to the user what needs to be done to finalize the plan in terms of establishing links that could not be found, resoving ambiguities, or confirming uncertain equivalences.

As an example, consider finding the female unemployment rate in Mycenae between 2000 and 2010. The system should be aware of the ambiguity, realize that such a statistics is meaningless for the ancient city and proceed with its modern namesake. It should seach in data portals to find a relevant dataset for Mycenae, New York, fail, and use a geographic names dataset to search for the higher-level administrative region, until data can be identified. The system should also try to ease other restrictions besides geographic coverage, and realize that there is no obvious preference between (a) female unemployment rates in New York state between 2000 and 2010; (a) county-level unemployment rates between 2000 and 2010; and (c) the county-level average female unemployment rate between 2005 and 2015. The system should then present all query plans, the datasets needed for each, the disambiguation decision it made, and the uncertainties and approximations of each plan.

Within this broad setting, the student will explore the application of a solver or planner to formulate plans for dataset linking and querying; and of a mechanism to select among these plans. The work is expected to focus on planning using resources manually prepared in advance, to avoid the implementation details of how to query actual on-line portals.